Volume 301, number 1, 10–14 FEBS 10905 April 1992

Purification, inhibitory properties and amino acid sequence of a new serine proteinase inhibitor from white mustard (*Sinapis alba* L.) seed

Enea Menegatti[a], Gabriella Tedeschi[b], Severino Ronchi[b], Fabrizio Bortolotti[a], Paolo Ascenzi[c], Richard M. Thomas[d], Martino Bolognesi[e] and Sandro Palmieri[f]

[a]*Dipartimento di Scienze Farmaceutiche, Università di Ferrara, Via Scandiana 21, 44100 Ferrara, Italy,* [b]*Istituto di Fisiologia Veterinaria, Università di Milano, Via Celoria 10, 20100 Milano, Italy,* [c]*Dipartimento di Scienza e Tecnologia del Farmaco, Università di Torino, Via Pietro Giuria 9, 10125 Torino, Italy,* [d]*Institut für Polymere, ETH Zentrum, Universitätstraße 6, CH-8092, Zürich, Switzerland,* [e]*Dipartimento di Genetica e Microbiologia, Sezione di Cristallografia, Università di Pavia, Via Abbiategrasso 207, 27100 Pavia, Italy* and [f]*Istituto Sperimentale per le Colture Industriali, MAF, Via di Corticella 133, 40129 Bologna, Italy*

Received 12 February 1992

A new serine proteinase inhibitor, mustard trypsin inhibitor 2 (MTI-2), has been isolated from white mustard (*Sinapis alba* L.) seed by affinity chromatography and reverse phase HPLC. The protein inhibits the catalytic activity of bovine β-trypsin and bovine α-chymotrypsin, with dissociation constants (K_i) of 1.6×10^{-10} M and 5.0×10^{-7} M, respectively, at pH 8.0 and 21°C, the stoichiometry of both proteinase–inhibitor complexes being 1:1. The amino acid sequence of MTI-2, which was determined following *S*-pyridylethylation, is comprised of 63 residues, corresponding to a molecular weight of about 7 kDa, and shows only extremely limited homology to other serine proteinase inhibitors.

Serine proteinase inhibitor; Amino acid sequence; Serine proteinase inhibition; White mustard seed

1. INTRODUCTION

Numerous peptides and proteins with the ability to inhibit the activity of serine proteinases are found throughout the living world [1]. A classification containing at least 10 different families of these molecules has been proposed on the basis of amino acid sequence homology, reactive site assignment and inhibitory mechanism [2]. In the plant kingdom, these inhibitors are found particularly in Graminaceae and Leguminosae seeds and in Solanaceae tubers and, in general, fall in the soybean Kunitz, Bowman–Birk and potato inhibitor families. Their prevailing role seems to be the control of endogenous proteinases during seed dormancy and protection against the proteolytic enzymes of many parasites and insects [3]. During previous work on the isolation of the high molecular weight serine proteinase inhibitor from white mustard seed [4], MTI-1, a smaller inhibitory protein with a molecular weight of about 7 kDa was detected, and the present paper describes the purification, primary structure and inhibitory properties of the latter molecule, which has been named mustard trypsin inhibitor 2 (MTI-2). As far as is known, this is the first inhibitor purified from a plant belonging to the Cruciferae to have had its sequence determined, and the results show it is markedly different to those of other trypsin inhibitors, preventing the assignment of the protein to any of the families within the current inhibitor classification.

Abbreviations: MTI-1, high molecular weight mustard trypsin inhibitor; MTI-2, low molecular weight mustard trypsin inhibitor; TPCK-trypsin, bovine trypsin treated with *N*-α-tosyl-L-phenylalanine chloromethyl ketone; β-trypsin, bovine β-trypsin; α-chymotrypsin, bovine α-chymotrypsin; BAPNA, *N*-α-benzoyl-L-arginine *p*-nitroanilide; ZTyrONp, *N*-α-carbobenzoxy-L-tyrosine *p*-nitrophenyl ester; RBTBBI, rice bran trypsin Bowman–Birk inhibitor.

Correspondence address: E. Menegatti, Dipartimento di Scienze Farmaceutiche, Università di Ferrara, Via Scandiana 21, 44100 Ferrara, Italy. Fax: (39) (532) 740142.

2. EXPERIMENTAL

2.1. *Materials*

The seeds of a commercial variety of white mustard (*Sinapis alba* L.) cv. Albatros, were purchased from the SIS Foraggera (Bologna, Italy). Sequencer reagents were obtained from Applied Biosystems (Foster City, CA, USA) and HPLC solvents from Merck AG (Darmstadt, Germany). Endoproteinases Lys-C and TPCK-trypsin were supplied by Boehringer AG (Mannheim, Germany), and bovine β-trypsin [5] (TRL, 3× crystallized, salt free) and bovine α-chymotrypsin for kinetic studies by the Worthington Chemical Co. (Freehold, NJ, USA). BAPNA, ZTyrONp and 4-vinylpyridine were from Sigma Chemical Co. (St. Louis, MO, USA) and all other reagents from Aldrich Chimica S.r.l. (Milan, Italy). All chemicals were reagent grade.

2.2. *Purification*

Mustard seed was homogenized in distilled water with a Ultra Turrax model T45 homogenizer (IKA-Werke, Stauffen, Germany), centrifuged and then treated as previously reported [4]. The crude extract was heated at 80°C for 3 min, centrifuged and the supernatant was concentrated by ultrafiltration on an Amicon YM-2 membrane

 Published by Elsevier Science Publishers B.V.

and then loaded onto a trypsin-Sepharose 4B column (5 × 25 cm) which had been equilibrated in 0.1 M triethanolamine, 0.1 M NaCl, 0.01 M $CaCl_2$, pH 7.8. Absorbed material was eluted with 0.3 M KCl, pH 1.8, and fractions showing high trypsin inhibitory activity were pooled, concentrated by ultrafiltration on an Amicon YM-2 membrane, and subjected to gel filtration on Sephadex G50SF (1 × 90 cm), using 0.05 M ammonium acetate buffer, pH 5.4, as eluent. Active fractions were pooled and further purified by RP-HPLC on a Vydac C-18 column (25 × 0.46 cm) using a linear gradient from 5 to 60% acetonitrile in 0.05 M KH_2PO_4, pH 6.0, at a flow rate of 1 $cm^3 \cdot min^{-1}$ at room temperature. The eluent was monitored by following UV-absorbance at 280 nm. HPLC separations were carried out with a Bruker model LC 41CD instrument equipped with a Chrom-A-Scope rapid scan detector (Bar Spec, Rehovot, Israel).

Molecular weights were determined by SDS-PAGE electrophoresis [6], and by gel filtration on Sephadex G50SF.

2.3. *Determination of dissociation constants*

Values of the dissociation constants (K_d) for the binding of MTI-2 to β-trypsin and α-chymotrypsin were determined from the results of activity measurements using BAPNA and ZTyrONp as substrates [5] at pH 8.0 and 21°C.

2.4. *Sequence determination*

2.4.1. *S*-pyridylethylation

MTI-2 (0.4 mg) was reduced with dithiothreitol (15 mM) in 400 μl of 0.25 M Tris–HCl, pH 8.5, for 4 h at 50°C and subsequently *S*-pyridylethylated with 4-vinylpyridine (0.3 mM) for 4 h at room temperature, both steps being performed under nitrogen and in the dark [7]. The protein was precipitated by the addition of 10% trichloroacetic acid, on ice, and the precipitate washed with a mixture of cold acetone and 1 M HCl (39:1).

2.4.2. Enzymatic cleavage

Tryptic cleavage was performed in 100 μl of 0.1 M *N*-ethylmorpholine acetate, pH 8.0, 37°C, for 3 h with an enzyme/substrate ratio of 1:25 (w/w). The digestion of *S*-pyridylethylated MTI-2 with endoproteinase Lys-C was carried out in 0.1 M Tris–HCl, 1 mM EDTA, pH 8.5, containing 10% (v/v) acetonitrile and 1 M urea in order to solubilize the material completely, at 37°C, for 20 h and with an enzyme/substrate ratio of 1:20. Enzyme digestions were stopped by the addition of TFA and the products separated by HPLC on an Aquapore RP-300 C8 column (25 × 0.46 cm). The solvent system used was: (A) 0.1% (v/v) TFA in water; (B) 0.075% (v/v) TFA in acetonitrile and peptides were eluted with a linear gradient from 0 to 50% B in A, in 70 min, at a flow rate of 0.7 $cm^3 \cdot min^{-1}$, UV-absorbance of the eluent being monitored at 220 nm. Peptide-containing fractions were concentrated under nitrogen and were used for amino acid composition and sequence analyses without further purification.

2.4.3. Amino acid analysis

Amino acid analysis was performed by the post-column *o*-phthalaldehyde derivatization procedure using a Jasco amino acid analyzer (Jasco, Tokyo, Japan) and following the method of Fujiara et al. [8]. Gas-phase hydrolysis was carried out in 6 M HCl, 1% (v/v) phenol, at 110°C for 24 h. Cysteine was estimated as cysteic acid following either (a) hydrolysis in the presence of dimethyl sulfoxide [9] of (b) oxidation of the protein with performic acid [10]. Tryptophan content was determined by second derivative spectroscopy of the protein in 6 M guanidine-HCl [11].

2.4.4. Amino acid sequence determination

Automated sequence analysis was performed on a pulsed-liquid sequencer mod. 477 (Applied Biosystems, Foster City, CA, USA) equipped with a 120A Applied Biosystems PTH-analyzer. Vapour phase modification of sulphydryl groups with 4-vinylpyridine for the N-terminal microsequence analysis of the entire protein was performed according to Amons [12].

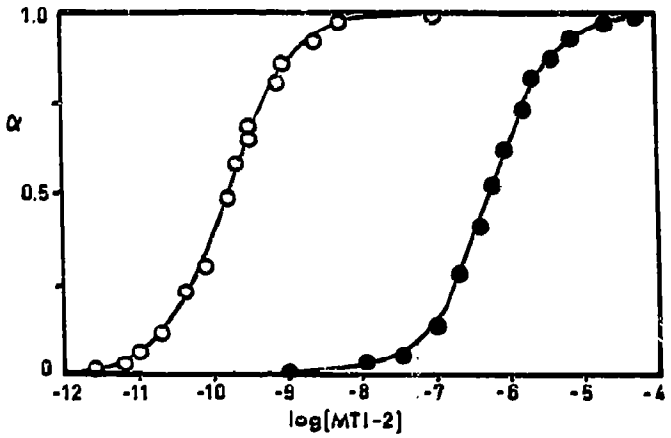

Fig. 1. The amount of MTI-2 bound to β-trypsin (○) and α-chymotrypsin (●) as a function of free inhibitor concentration (log[MTI-2]). Solid lines were generated by the equation:

$$\alpha = \frac{1}{(1 + 10^{[\log_{10} K_d)/(10^{\log_{10}[MTI\text{-}2]})}}$$

using $K_d = 1.6 \times 10^{-10}$ M and $K_d = 5.0 \times 10^{-7}$ M for β-trypsin:MTI-2 and α-chymotrypsin:MTI-2 complex formation, respectively, and where α is the fraction of proteinase molecules with bound inhibitor. Values of K_d were obtained with an iterative non-linear least-squares curve-fitting procedure. A standard deviation of ±8% was calculated for the K_d values from the fitting procedure. Data were obtained at pH 8.0 (Tris–HCl, 0.1 M) and 21°C. For further details, see text.

2.4.5. Computer sequence analysis

Homologies with the entries in the Swiss-Prot Database were searched for by the PC-GENE program [13].

2.5. *Peptide nomenclature*

Peptides were numbered according to their position in the MTI-2 sequence, the N-terminal residue of the molecule being assigned the number 1. Tryptic peptides are indicated by the prefix T-, endoproteinase Lys-C peptides by the prefix L- and N-terminal sequence by the prefix N-.

3. RESULTS

3.1. *Purification and inhibitory properties*

Crude mustard seed extract possesses both anti-trypsin and anti-chymotrypsin activities and, as has been reported previously [4], a trypsin inhibitor, MTI-1, which strongly inhibits bovine β-trypsin but exhibits no activity towards bovine α-chymotrypsin, can be isolated from the extract using a combination of affinity and ion-exchange chromatographies. MTI-1 is rapidly and irreversibly deactivated on heating to 80°C. However, when the whole soluble mustard seed protein extract is subjected to such treatment, anti-serine proteinase activity persists, indicating that there must be at least two inhibitors present in the extract. Taking advantage of the thermal lability of MTI-1 by incorporating a heating step into the purification procedure has facilitated the isolation of a second inhibitor, MTI-2, which has now

been purified in high yield. The protein has a molecular weight of 7.0±0.5 kDa, as judged by SDS-PAGE and gel filtration chromatography, and inhibits the activities of both β-trypsin and α-chymotrypsin. Fig. 1 shows the isotherms for MTI-2 binding to β-trypsin and α-chymotrypsin, using BAPNA and ZTyrONp as substrates, respectively. The dissociation constant for the β-trypsin:MTI-2 complex was found to be 1.6×10^{-10} M while that for the α-chymotrypsin:MTI-2 complex was 5.0×10^{-7} M (pH 8.0, 21°C) and the stoichiometry for complex formation was 1:1 in both cases. Values of K_d were independent of the enzyme and substrate concentrations, as expected for simple systems, the Hill coefficient being 1.00±0.02.

3.2. *Structural studies*

Table 1 gives the amino acid composition of MTI-2. No methionine or histidine are present and there is only one tryptophan residue, compositional features typical of plant serine proteinase inhibitors. The protein, following modification of thiol groups by pyridylethylation, was subjected to N-terminal sequence analysis and 39 Edman cycles were identified unambiguously, as is shown in Fig. 2. The sample appeared to be homogeneous, as indicated by the presence of a single amino acid residue at each degradation cycle.

Reduced and *S*-pyridylethylated MTI-2 was digested with TPCK-trypsin and the resultant peptides were separated by HPLC (Fig. 3). Eleven major peptides were isolated in a pure form and were sequenced (Fig.2).

Table 1

Amino acid composition of MTI-2

Amino acid	Mol. residue/mol. protein	Sequence data
Asx	7.8 (8)	8
Thr	2.9 (3)	3
Ser	2.1 (2)	2
Glx	4.8 (5)	5
Pro	4.0 (4)	4
Gly	7.8 (8)	8
Ala	2.4 (2)	2
Cys[a]	7.0 (7)	8
Val	1.6 (2)	2
Met	0.0 (0)	0
Ile	2.1 (2)	3
Leu	2.0 (2)	2
Tyr	3.0 (3)	3
Phe	3.7 (4)	4
Lys	3.8 (4)	4
His	0.0 (0)	0
Arg	4.0 (4)	4
Trp[b]	1.0 (1)	1

Values are given as residues per molecule with the nearest whole number in parentheses. Values derived from the sequence determination are given in the second column.
[a]Determined as cysteic acid
[b]Determined by second derivative spectroscopy

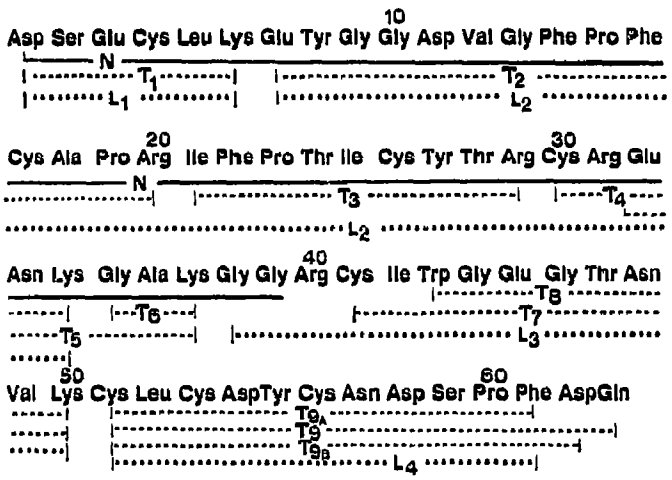

Fig. 2. The primary structure of MTI-2. The prefixes T, L or N indicate sequences that were derived from tryptic or Lys-C cleavage or from N-terminal sequencing, respectively. Automated Edman degradation was performed on 300–500 pmol aliquots of each peptide and on about 1 nmol of the entire protein. In all cases the average repetitive yield was 94%.

Peptide T9 was tentatively assigned to the C-terminus since it contained neither lysine nor arginine. All tryptic peptides analyzed originated from cleavage at expected sites with the exception of T8 which was produced by the abnormal hydrolysis of the Ile^{42}–Trp^{43} peptide bond. Reduced and *S*-pyridylethylated MTI-2 was also digested with endoproteinase Lys-C in order to produce sequence overlaps with the peptides derived from tryptic digestion. Four major peptides were isolated and sequenced (Figs. 2 and 4). Peptide L4 was assigned to the C-terminus as it lacked a terminal lysine residue.

Fig. 2 shows the complete sequence of MTI-2. The calculated molecular weight of 7.036 kDa is in excellent

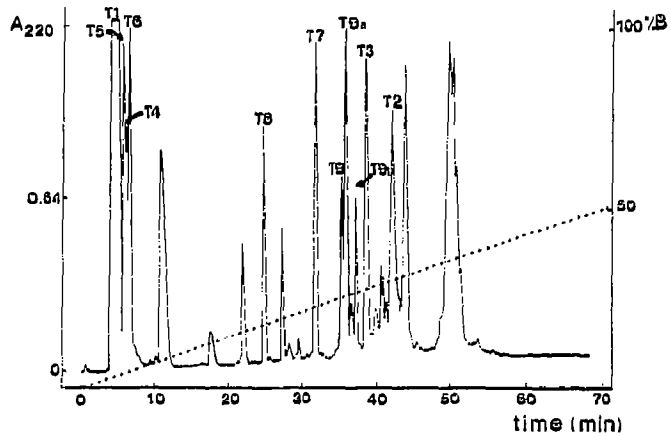

Fig. 3. Reverse phase HPLC purification of the tryptic peptides. Peptides were separated on an Aquapore RP-300 C_8 column with a linear gradient from 0–50% B in 70 min. The solvents used were: (A) 0.1% TFA in water and (B) 0.075% TFA in acetonitrile. The effluent was monitored at 220 nm.

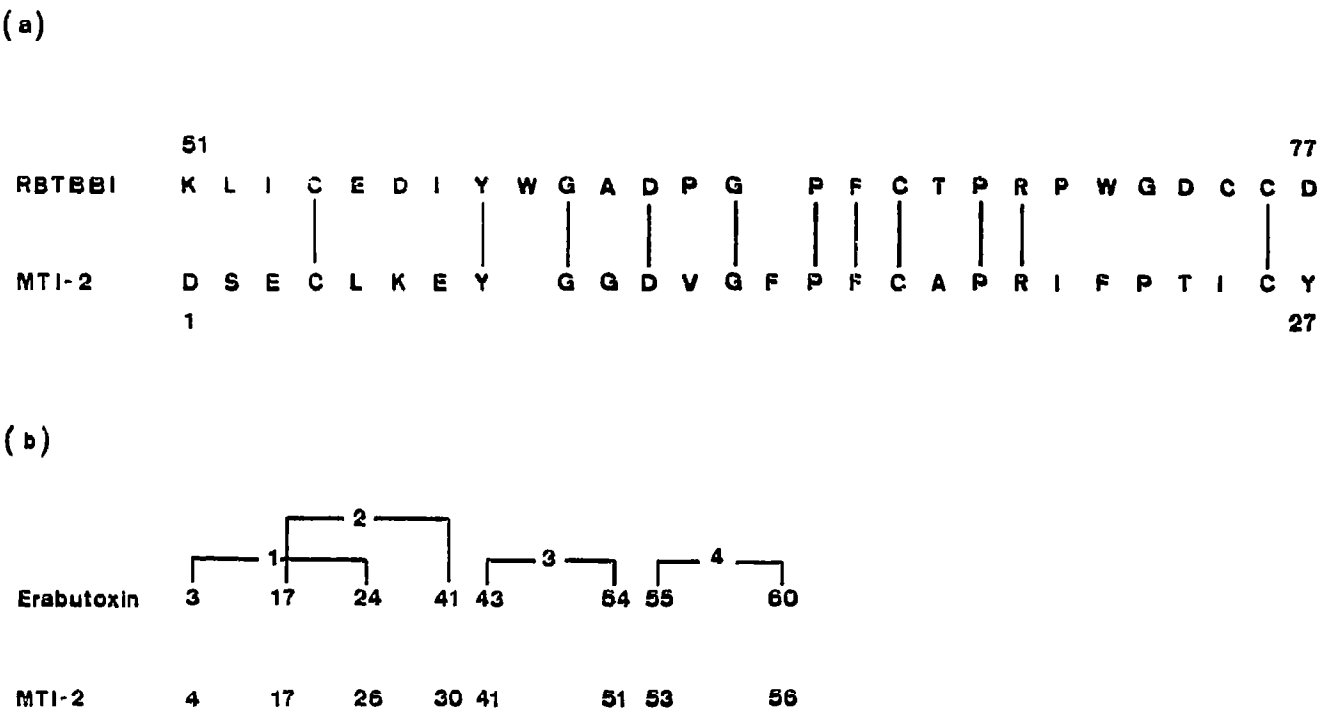

Fig. 4. (a) Sequence homology between the primary sequences of MTI-2 (residues 1–27) and rice bran trypsin inhibitor (RBTBBI) (residues 51–77). Identical residues are shown by continuous lines. (b) The position of cysteine residues in MTI-2 and erabutoxin. Disulphide bond positions in erabutoxin are also shown.

agreement with the molecular weight determined by gel filtration and SDS-PAGE (7.0±0.5 kDa)

In some cases different batches of the inhibitor exhibited sequence microheterogeneity. Species lacking the N-terminal residue, Asp^1 and variants terminating at Phe^{61}, Asp^{62} as well as Gln^{63} have been detected.The microheterogeneity probably arises as a result of proteolysis during either seed dormancy or the purification procedure.

4. DISCUSSION

MTI-2 is the second serine proteinase inhibitor to have been isolated from white mustard seed but is the first to have had its amino acid sequence determined. MTI-2 differs considerably from MTI-1 and the possibility that it arises from the cleavage of MTI-1 can be ruled out because (i) there are 8 cysteine residues in MTI-2 but only 4 in MTI-1 [4], (ii) MTI-2 is thermally stable whereas MTI-1 is not, and (iii) MTI-2 strongly inhibits both β-trypsin and αchymotrypsin, while MTI-1 shows no activity towards α-chymotrypsin [4]. The K_d values obtained for the serine proteinase:MTI-2 complexes indicate that they are remarkably stable and that the inhibitor possesses a higher affinity for β-trypsin than MTI-1.

The complete amino acid sequence of MTI-2 has been determined (Fig.2) and consists of 63 residues. The sequence data are in good agreement with the amino acid composition reported in Table 1. The only difference appears to be in the Cys content. This last value is, however, within the recovery limits of the gas-phase hydrolysis method used [14]. The absence of free thiol groups determined by the *p*-chloromercuribenzoate procedure [15] strongly supports the presence of 4 disulphide bridges in the molecule.

The amino acid sequence of the protein is considerably different to those of other trypsin inhibitors and, as it cannot be included within the current classification of these proteins, it may belong to an entirely new inhibitor family. Only one region of the sequence was found to have any homology with another inhibitor: residues 8–21 showed some similarity to a portion of the second domain of RBTBBI [16] (Fig. 4a), a region that does not, however, include the reactive site of the protein, but only the connecting peptide between domains II and III. No direct experimental evidence is yet available concerning the reactive site of MTI-2 although studies aimed at its identification are in progress.

The presence of 8 cysteine residues on the sequence suggested possible similarities between MTI-2 and proteins whose structure is organized around a four- disulfide core. This folding pattern is thought to stabilize proteins or domains too small to have a significant hydrophobic core, like snake venom neurotoxins, cardiotoxins, cytotoxins, hevein, wheat germ agglutinin and ragweed pollen allergen as well as anti-leukoproteinases and basic proteinase inhibitors. According to

Drenth et al. [17], many of these molecules are classified as belonging to this group on the basis of the cysteine alignment alone, although, in some cases, no sequence homology or functional similarities have been found. As shown in Fig. 4b, the positions of the cysteine residues in MTI-2 are similar to those found in erabutoxin although considerably more structural data will be required in order to confirm whether an agglutinin/ neurotoxin-like fold is indeed present in MTI-2.

REFERENCES

[1] Laskowski, M. and Laskowski Jr., M. (1954) Adv. Protein Chem. 9, 203–242.
[2] Laskowski Jr., M. and Kato, I. (1980) Annu. Rev. Biochem. 49, 593–626.
[3] Ryan, C.A. (1981) in: The Biochemistry of Plants, vol. 9, pp. 351–370, Academic Press, New York.
[4] Menegatti, E., Palmieri, S., Walde, P. and Luisi, P.L. (1985) J. Agric. Food Chem. 33, 784–789.
[5] Menegatti, E., Boggian, M., Ascenzi, P. and Luisi, P.L. (1987) J. Enzyme Inhib. 2, 67–71.
[6] Laemmli, U.K. (1970) Nature 227, 680–685.
[7] Friedman, M., Krull, L.H. and Cavins, J.F. (1970) J. Biol. Chem. 245, 3868–3871.
[8] Fujiara, M., Ishida, Y., Nimura, N., Toyama, A. and Kinoshita, T. (1987) Anal. Biochem. 166, 72–78.
[9] Spencer, R.L. and Wold, F. (1969) Anal. Biochem. 32, 185–190.
[10] Moore, S. (1963) J. Biol. Chem. 238, 235–237.
[11] Nozaki, Y. (1987) Arch. Biochem. Biophys. 277, 324–333.
[12] Amons, R. (1987) FEBS Lett. 212, 68–72.
[13] Bairoch, A. and Boeckmann, B. (1991) Nucleic Acids Res. 19, 2247–2249.
[14] Spencer, R.L. and Wold, F. (1969) Anal. Biochem. 32, 185–190.
[15] Boyer, P.D. (1954) J. Am. Chem. Soc. 76, 4331–4337.
[16] Tashiro, M., Hashino, K., Shiozaki, M., Ibuki, F. and Maki, Z. (1987) J. Biochem. 102, 297–306.
[17] Drenth, J., Low, B.W., Richardson, J.S. and Wright, C.S. (1980) J. Biol. Chem. 255, 2652–2655.
[18] Antonini, E., Ascenzi, P., Bolognesi, M., Gatti, G., Guarneri, M. and Menegatti, E. (1983) J. Mol. Biol. 165, 543–558.